The Journal of Financial Research

RAWLS COLLEGE OF BUSINESS, TEXAS TECH UNIVERSITY
PUBLISHED FOR THE SOUTHERN AND SOUTHWESTERN
FINANCE ASSOCIATIONS BY WILEY-BLACKWELL PUBLISHING

The Journal of Financial Research • Vol. XXXVII, No. 4 • Pages 543–551 • Winter 2014

WHAT DOES $\beta_{SMB} > 0$ REALLY MEAN?

Hsiu-lang Chen and Gilbert Bassett

University of Illinois at Chicago

Abstract

A positive SMB coefficient in a Fama–French regression is often interpreted as signaling a portfolio weighted toward small-cap stocks. We present a very large portfolio, which has a positive SMB coefficient for all periods. We emphasize that this is associated with the coexistence of both "M"—the market—and "SMB"—the mimicking portfolio for size—in the Fama–French three-factor model. We explain why the model can attribute small size to large-cap stocks and portfolios. The results highlight how coefficients should be interpreted when a self-financing portfolio is used for portfolio attribution.

JEL Classification: G10, G11

I. Introduction

The Fama–French three-factor model has become the standard academic tool for assessing portfolios as well as individual stocks. The three factors are: (1) a market factor —RMRF, (2) a size factor—SMB, and (3) a value factor—HML. The model is often used to identify exposure to the factors—the portfolio's "style."[1] Factor investing has recently gained attention from the financial press and has been finding favor among institutional investors and high-end financial advisers.[2] As such, it is essential to understand the meaning of such attribution and particularly the way the inclusion of mimicking portfolios might affect the interpretation of regression loadings.

The coefficients in the Fama–French regression are often interpreted in absolute terms, so that, for example, a positive SMB coefficient would indicate a portfolio that favors small-cap stocks. A recent analysis of a universe of mutual funds, for example, concluded that there was a general tendency for the funds to hold small stocks because

We are grateful to the referee and Harry Turtle (the associate editor) for their comments. We also thank seminar participants at University of Illinois at Chicago and the annual meeting of the Global Finance Conference, Chicago,May 23–25, 2012. This article supersedes our working paper previously circulated as "Returns-Based Attribution with Fama–French Factors."

[1]Returns-based attribution uses time-series returns of a portfolio with unknown constituents to derive estimates of the portfolio's "factor" exposures. The regression coefficients on the returns of factor-based portfolios provide estimates of the portfolio's factor exposure. The constant term shows the portfolio's expected return after controlling for a passive portfolio invested in the regression-weighted factors.

[2]J. Light and B. Levisohn, "Here's What's Really Driving Your Returns," *Wall Street Journal* (December 24, 2011), B5.

the average SMB coefficient of the funds in the universe was a positive 0.1628.[3] These conclusions about the relation between a positive SMB coefficient and small stocks, however, are incorrect: a positive SMB coefficient does not necessarily mean returns are attributable to exposure to small stocks.

The cleanest way to see that this occurs is to consider a reference portfolio whose constituents are known and identical to the "S" and "B" returns in SMB. Our reference portfolio is tilted toward large-cap stocks by virtue of being 80% weighted on B returns and 20% on S. In spite of the 80% weight on B, the three-factor attribution always gives a positive SMB coefficient. The seeming contradiction occurs for two reasons. The first is that big firms account for most of the market value in the stock market. The second is related to the self-financing SMB portfolio that is included in the model. Figure I shows that the small portfolio (S) is composed of about 80% of all publicly traded firms, but this represents less than 10% of market value. As a result, the 80% large-cap portfolio is in fact small when compared to the overall market. If, instead of the Fama–French model, we apply an attribution model such as proposed by Sharpe (1992), in which only non-self-financing portfolios are used—that is, separate S and B regressors replace SMB—then the attribution correctly indicates the portfolio composition.[4]

Self-financing portfolios are commonly used in the theory and practice of finance. However, potential issues of combining self-financing portfolios and standard portfolios have not been fully explored. Korkie and Turtle (2002) quantify the impact on the efficient frontier when a self-financing portfolio is added to a standard portfolio with weights summing to one. They show that the Fama–French three factors do not fully span the entire asset universe. Our study addresses the issue of coefficient interpretation on a self-financing portfolio in the portfolio attribution, which has not been explored in the literature. Thus, our study contributes to the literature in two ways. First, we document a positive SMB coefficient for large portfolios and individual stocks. Second, we provide explanations for how the SMB coefficient should be interpreted.

[3] Elton, Gruber, and Blake (2011, p. 348) state, "When we examine the small-minus-big factor, we see that the average beta is 0.1628, demonstrating a general tendency for funds to hold small stocks. However, over 25% of our funds have a negative beta with the size factor, which indicates that they are overweight large stocks." As another example, Carhart (1997 p. 61) uses the model with a fourth momentum factor "as a performance attribution model, where the coefficients and premia on the factor-mimicking portfolios indicate the proportion of mean return attributable to four elementary strategies: high versus low beta stocks, large versus small market capitalization stocks, value versus growth stocks, and one-year return momentum versus contrarian stocks." For a final example, Fama and French (2010, p. 1944) say, "For example, consider an actively managed small value fund. The passive benchmark for the fund produced by the three-factor version of (1) [the model] is likely to imply positive weights on the market, SMB, and HML, which implies positive weight on the market(M), small stocks (S), and value stocks (H) and negative weights on big stocks (B) and growth stocks (L)."

[4] We perform the Sharpe asset allocation model and Fama–French risk factor model on the reference portfolio every year since 1926. In the Sharpe model, we try two sets for the independent variables. One includes B and S only, whereas the other includes the market, B, and S. In a modified Fama–French model, RMRF and SMB are the independent variables. The result shows that the Sharpe model correctly identifies the percentage weights of the reference portfolio in every test year whereas the modified Fama–French model only identifies a negative coefficient on SMB in 8 of 84 test years. The result is available upon request.

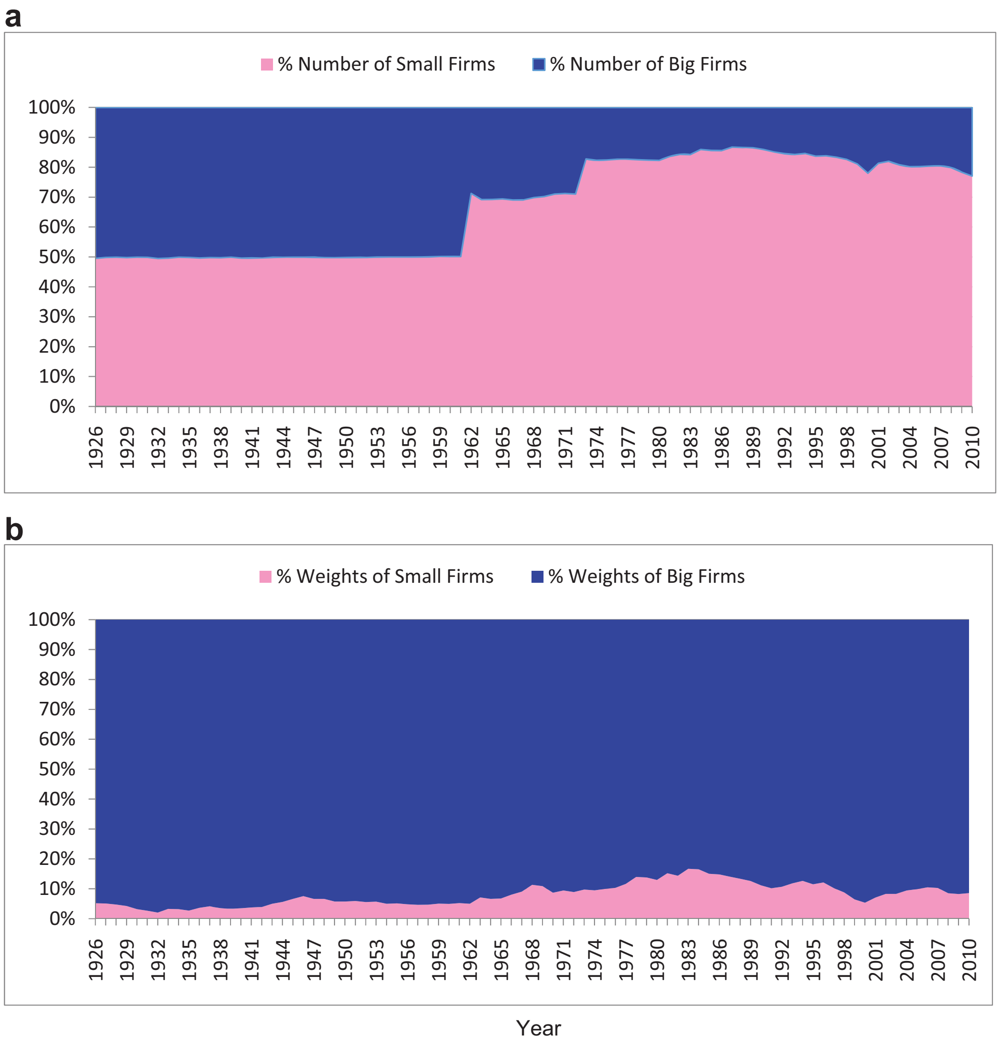

Figure I. Size Composition of the Stock Market. We construct the median breakpoint for size based on the market capitalization of all firms listed on the NYSE at the end of each June since 1926. A firm in NYSE/AMEX/NASDAQ is assigned to the small firm and big firm portfolios accordingly. We only include common stocks with Center for Research in Securities Prices (CRSP) share codes 10 and 11. Note that AMEX stocks are introduced beginning July 1962 and NASDAQ stocks are introduced beginning December 14, 1972. Figure Ia shows the number of firms in the market by percentage, and Figure Ib shows the percentage by market capitalization.

II. Data

The Center for Research in Security Prices (CRSP) return files and Kenneth French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html) constitute our main data sources. We follow Fama–French methodology to construct individual constituents S and B in the SMB factor for reference portfolios. Our sample includes common stocks with CRSP share codes of 10 or 11 from June 1926 to May 2011.

TABLE 1. Sample Statistics for Portfolio 8B +.2S.

	Portfolio	RMRF	SMB	HML	SMALL	BIG	HIGH	LOW
				Average				
	1.07	0.63	0.25	0.39	**1.26**	**1.02**	1.34	0.95
				Correlation				
Portfolio	1.00	0.98	**0.39**	0.39	0.95	0.99	0.97	0.95
RMRF		1.00	0.33	0.23	0.91	0.98	0.90	0.96
SMB			1.00	0.10	0.66	0.29	0.49	0.53
HML				1.00	0.35	0.40	0.57	0.12
SMALL					1.00	0.91	0.96	0.95
BIG						1.00	0.95	0.92
HIGH							1.00	0.89
LOW								1.00
				Covariance				
Portfolio	37.75	32.72	7.90	8.65	44.07	36.17	45.49	36.84
RMRF		29.73	5.93	4.43	37.46	31.53	37.63	33.20
SMB			11.04	1.16	16.73	5.70	12.40	11.24
HML				12.79	9.58	8.42	15.49	2.70
SMALL					57.46	40.72	55.41	45.83
BIG						35.03	43.01	34.59
HIGH							58.41	42.92
LOW								40.23

Note: Monthly returns from July 1926 to May 2011 of six research portfolios—small growth (SG), small neutral (SN), small value (SV), big growth (BG), big neutral (BN), and big value (BV)—used for Fama and French (1996) three-factor construction as well as the three-factors, RMRF, SMB, and HML (as provided by French). Statistics are for a portfolio allocating 80% in BIG and 20% in SMALL. BIG is the B in SMB, namely, the average of BG, BN, and BV, and SMALL is the S in SMB, namely, the average of SG, SN, and SV. HIGH is the simple average of SV and BV, and LOW is the simple average of SG and BG. The monthly sample mean return is in a percentage format and return covariance in a 10^{4} format.

III. SMB Attribution in the Three-Factor Model

We consider a reference portfolio consisting of 80% big (B) and 20% small (S) where B and S are the same as in SMB; that is, returns are .8B +.2S. Table 1 presents sample statistics for this portfolio. It shows, for example, that the average monthly return for small stocks (1.26) has exceeded the return (1.02) to big stocks. But already anticipating our findings, notice that in spite of the 80% weight on B, the portfolio's correlation with SMB is a positive 0.39.

Three-factor estimates are shown in Table 2. Panel A presents the three-factor regressions for the reference portfolio as well as portfolios in which B and S vary in increments of 10%; that is, returns are $\lambda B + (1 - \lambda)S, \lambda = 0, 0.1, 0.2, \ldots, 0.9, 1.0$. The table shows that (except when $S = 0\%$) the SMB coefficient is positive. The same pattern shows up in Panel B where the regression excludes HML. Note that the non-SMB coefficients for the different portfolios are all identical. This is a direct consequence of

TABLE 2. Three-Factor Attribution for Portfolios Based on Convex Combinations of S and B.

Reference Portfolio	Constant	RMRF	SMB	HML	Adj R^2
Panel A. Three-Factor Estimations					
100%BIG, 0%SMALL	−0.03 [−1.70]	1.03 [285.82]	**−0.07 [−11.29]**	0.31 [58.59]	99.04
90%BIG, 10%SMALL	−0.03 [−1.69]	1.03 [285.82]	**0.03 [5.98]**	0.31 [58.59]	99.08
80%BIG, 20%SMALL	**−0.03 [−1.69]**	**1.03 [285.81]**	**0.13 [23.24]**	0.31 [58.59]	**99.11**
20%BIG, 80%SMALL	−0.03 [−1.69]	1.03 [285.78]	**0.73 [126.80]**	0.31 [58.59]	99.35
10%BIG, 90%SMALL	−0.03 [−1.69]	1.03 [285.78]	**0.83 [144.06]**	0.31 [58.59]	99.38
0%BIG, 100%SMALL	−0.03 [−1.69]	1.03 [285.77]	**0.93 [161.31]**	0.31 [58.59]	99.42
Panel B. Bivariate Estimation Excluding HML					
100%BIG, 0%SMALL	0.06 [1.52]	1.07 [145.59]	**−0.06 [−4.68]**		95.81
90%BIG, 10%SMALL	0.06 [1.52]	1.07 [145.58]	**0.04 [3.57]**		95.95
80%BIG, 20%SMALL	**0.06 [1.52]**	**1.07 [145.58]**	**0.14 [11.82]**		96.11
20%BIG, 80%SMALL	0.06 [1.52]	1.07 [145.57]	**0.74 [61.33]**		97.13
10%BIG, 90%SMALL	0.06 [1.52]	1.07 [145.57]	**0.84 [69.58]**		97.29
0%BIG, 100%SMALL	0.06 [1.52]	1.07 [145.57]	**0.94 [77.83]**		97.45

Note: Coefficient estimates are reported for regressions using monthly returns from July 1926 to May 2011; t-statistics are reported in brackets. The reference portfolio is a combination of BIG and SMALL where BIG and SMALL are the constituents of SMB. The results are for selected reference portfolios. Panel A shows the results for Fama–French three-factor estimations. Panel B shows the results for bivariate (no HML) estimations.

Fama–French three-factor model:

$$R_{\text{Portfolio}} - R_f = \alpha + \beta_1 \text{RMRF} + \beta_2 \text{SMB} + \beta_3 \text{HML} + \varepsilon.$$

the linearity of the least squares estimator when a self-financing portfolio is included in the independent variable set to explain a portfolio returns.[5] Another consequence is that the SMB coefficient is a simple function of the λ-weight in the portfolio, namely, $\beta_{SMB}(\lambda) = \beta_{SMB}(0) - \lambda$.

For the estimates in Table 2, $\beta_{SMB}(1) = -.07, \beta_{SMB}(0) = .93$ so $\beta_{SMB}(\lambda) = -.07\lambda + .93(1 - \lambda)$. All it takes for our $\lambda B + (1 - \lambda)S$ portfolio to register "small" is that it has more than $1 - .93 = .07$ weight on small returns. This estimated cutoff weight is very close to the actual percentage weight of big firms in the market as shown in Figure I. The big firms (B) account for about 92% of total market value on average. The positive SMB coefficient for an 80% large-cap portfolio occurs because the portfolio is in fact "small," at least in comparison to the overall market.[6]

[5] That is, partition X into $X = [i|Z|S-B]$, where i is an n-vector of ones, Z is an $n \times k$ submatrix of explanatory variables, and $S-B$ is the n-vector of small-minus-big returns. The y-vector of dependent variables is $y = \lambda B + (1-\lambda)S$. Write the partitioned least squares vector, which depends on λ, as $\beta(\lambda) = [\beta_0(\lambda)|\beta_Z(\lambda)|\beta_{SMB}(\lambda)]$. Noting that $[S-B] = X\begin{bmatrix}0\\0\\1\end{bmatrix}$, so $X'[S-B] = X'X\begin{bmatrix}0\\0\\1\end{bmatrix}$ and the least squares estimate is $\beta(\lambda) = X'X^{-1}X'S - \lambda \begin{bmatrix}0\\0\\1\end{bmatrix} = [\beta_0(0)|\beta_Z(0)|\beta_{SMB}(0) - \lambda]$.

[6] The same relative interpretation also holds for the HML coefficient. The H and L components in HML, however, have varied over time relative to the overall market so that unlike SMB there has not been a consistent tendency for HML coefficients to be positive for L-dominated portfolios.

It should be noted that although the Fama–French model is often used for attribution, it grew out of attempts to understand anomalies relative to the one-factor market model. In contrast, Sharpe (1992) attribution excludes the market and uses mutually exclusive non-self-financing-style portfolios (e.g., large/value, large/growth, small/value, small/growth), and thus, it will not have the issue—an error inference on the portfolio attribution; for an application of Sharpe-style attribution, see Bassett and Chen (2001).

The tendency for a positive factor loading on SMB has been noted at the portfolio level elsewhere. Chan, Chen, and Lakonishok (2002) show that less than 30% of mutual funds have a negative SMB coefficient even though more than 38% of mutual funds have a weighted-average size rank that is larger than the median of S&P 500 firms. In a similar vein, Daniel and Titman (1997, p. 6) indicate from their analysis of returns that "after controlling for book-to-market there is more of a large firm rather than a small firm anomaly." In addition, Fama and French (1993, 1996) show that only the largest quintile size portfolio has a negative factor loading on SMB. Whether the tendency for a positive factor loading on SMB still exists for individual firms has not been noted in the literature. We examine this next.

IV. SMB Attribution for Individual Stocks

Attribution for individual stocks is reported in Table 3. As in Fama and French (1996) decile breakpoints for size-based capitalization of NYSE firms are constructed at the end of each June. Each firm in NYSE/AMEX/NASDAQ is then assigned to the decile portfolio accordingly. At the beginning of every July, the Fama–French three-factor regression is performed using monthly returns over the following year. The dependent variable is a stock's (excess to the T-bill rate) return. For each decile, the table shows (1) the number of stocks, (2) the average β_{SMB}, (3) the proportion of stocks with positive β_{SMB}, and (4) the proportion of stocks with a statistically significant positive β_{SMB}.

The individual stocks in the large-size deciles from 6 to 10 are—by construction—tilted to B whereas stocks in deciles 1 to 5 are tilted to S. Table 3 shows, however, that more than half of the stocks in deciles 1 to 9 have a positive exposure to SMB. On average, there were 125 firms in size decile 9, the average SMB coefficient was 0.09, and 52% had a positive SMB coefficient. Again, this tendency for large-cap stocks to look small is not a quirk of the returns in either subperiod. In the hypothesis test, although we can differentiate the average of SMB coefficients in two subperiods for the extreme deciles, we cannot differentiate it across 10 deciles.

For a final look at attribution in the three-factor model we consider annual attribution only for the individual stocks in the B portfolio. As depicted in Figure II, the proportion of B stocks with the wrong sign has been about 55%. The results based on individual firms again highlight the tendency for a positive factor loading on SMB to be greater than expected.

V. Univariate SMB Attribution

Although the market factor in the three-factor model is the primary reason large-cap portfolios have positive SMB coefficients, we might uncover further reasons for the

TABLE 3. Three-Factor SMB Attribution for Individual Stock by Size Deciles.

	Size Deciles									
Periods for Portfolio Formation	(Small) 1	2	3	4	5	6	7	8	9	(Big) 10
Whole Period: 192606~200906										
No. of firms	1,231	344	238	194	168	149	137	131	125	121
β_{SMB}	1.44	1.13	0.99	0.86	0.74	0.54	0.38	0.25	0.09	−0.15
% $\beta_{SMB} > 0$	0.73	0.73	0.73	0.71	0.69	0.65	0.61	0.57	0.52	0.42
% $\beta_{SMB}^* > 0$	0.15	0.15	0.15	0.14	0.13	0.11	0.09	0.07	0.06	0.04
Subperiod: 192606~196706										
No. of firms	169	98	92	91	89	89	89	88	89	90
β_{SMB}	1.67	1.22	1.04	0.89	0.77	0.55	0.4	0.28	0.13	−0.09
% $\beta_{SMB} > 0$	0.77	0.75	0.75	0.73	0.71	0.66	0.63	0.59	0.54	0.45
% $\beta_{SMB}^* > 0$	0.17	0.17	0.16	0.16	0.14	0.12	0.1	0.08	0.07	0.04
Subperiod: 196806~200906										
No. of firms	2,293	589	384	298	247	209	185	174	160	153
β_{SMB}	1.21	1.05	0.94	0.83	0.71	0.53	0.36	0.22	0.05	−0.21
% $\beta_{SMB} > 0$	0.69	0.71	0.71	0.7	0.68	0.64	0.6	0.56	0.51	0.4
% $\beta_{SMB}^* > 0$	0.13	0.14	0.14	0.13	0.11	0.09	0.08	0.07	0.06	0.03
Test H_0: β_{SMB} is identical in two subperiods										
Difference in β_{SMB}	0.46	0.17	0.1	0.05	0.06	0.02	0.03	0.06	0.08	0.13
p-value	.00	.01	.02	.16	.18	.71	.45	.22	.07	.00
Test H_0: β_{SMB} is identical in two subperiods across 10 deciles										
$F(1, 18) = 0.27$						p-value $= .77$				

Note: Decile breakpoints for size are based on the market capitalization of all firms listed on the NYSE at the end of each June since 1926. A firm in NYSE/AMEX/NASDAQ is assigned to the decile portfolio accordingly. (We include only common stocks with share codes 10 and 11.) At the beginning of every July, for each stock we run the Fama–French three-factor model based on monthly returns over the following one year. The dependent variable is a stock's return excess of the T-bill rate. The simple average of coefficient estimates is reported for each decile. To run a regression, we require a stock to have complete one-year return data following portfolio formation. The number of stocks by percentage having a positive SMB coefficient is reported in the row labeled "% $\beta_{SMB} > 0$," and the number of stocks by percentage having a significant and positive SMB coefficient at the 5% level is reported in the row labeled "% $\beta_{SMB}^* > 0$." Snapshot statistics are constructed for each year and the statistics averaged across time are presented for the whole period and two subperiods. We test whether the average of coefficient estimates is identical in two subperiods for each decile and across 10 deciles.

positive SMB coefficients by considering a simpler case in which the regression includes only SMB in the univariate regression.

In an unreported result, we run a univariate regression for the portfolio with returns, $\lambda B + (1 - \lambda)S$, where λ is in [0,1]. It shows that the coefficient on SMB, which we denote by β_{S-B}, is positive for all λ. To understand how this occurs, write the univariate coefficients in terms of the weights that define the λ-reference portfolios. With $P = \lambda B + (1 - \lambda)S$, we have

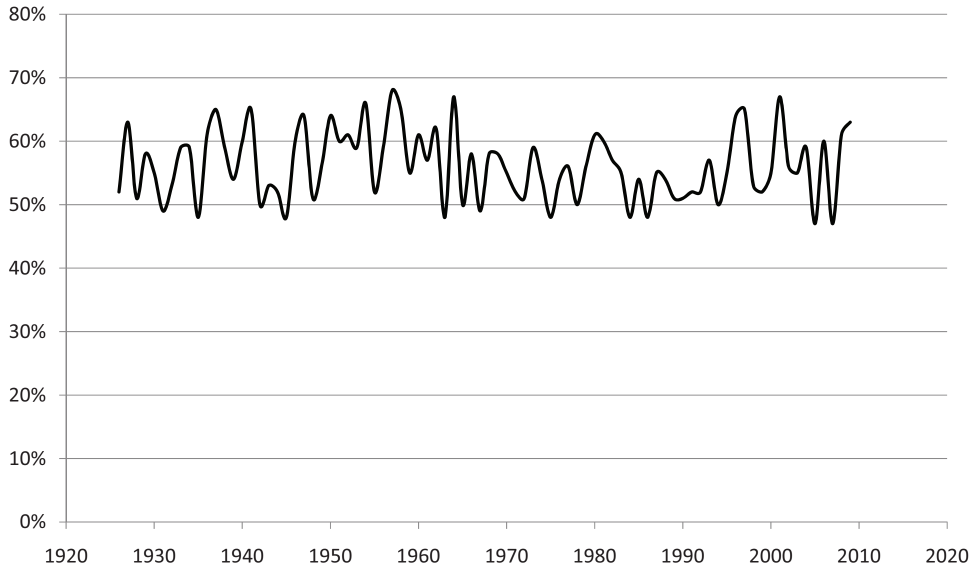

Figure II. Annual Percentage Big Stocks with $\beta_{SMB} > 0$.

$$\beta_S = \frac{Cov((1-\lambda)S + \lambda B, S)}{Var(S)} = (1-\lambda) + \lambda\frac{Cov(B,S)}{Var(S)}$$

$$\beta_B = \frac{Cov((1-\lambda)S + \lambda B, B)}{Var(B)} = \lambda + (1-\lambda)\frac{Cov(B,S)}{Var(B)}$$

and

$$\beta_{S-B} = (1-\lambda) + \frac{Cov(B,S) - Var(B)}{Var(S-B)}.$$

Substituting the values of the variances and covariances from Table 1, we find that even at the extreme, $\lambda = 1$—the entire portfolio is B—the coefficient β_{S-B} is positive. The coefficient says "small" even though the entire portfolio is "big."

Alternatively, the coefficient of $S-B$ can be expressed in terms of the β_B and β_S coefficients on the separate B and S variables, or

$$\beta_{S-B} = \frac{Cov(P,S) - Cov(P,B)}{Var(S-B)} = \frac{\beta_S Var(S) - \beta_B Var(B)}{Var(S-B)}.$$

If returns are more highly correlated with B than S, one might expect $\beta_{S-B} < 0$, and this is valid when Var(B) = Var(S). But when Var(S) > Var(B), which is typically the case, it becomes possible for $\beta_{S-B} > 0$ even though $\beta_S - \beta_B < 0$. The condition for this is,

$$\beta_B < \beta_S \frac{Var(S)}{Var(B)}.$$

Inspection of β_{S-B}, β_S, and β_B, we find that this condition is satisfied in every referenced portfolio.

Another take on the same finding, which is not reported, mimics the previous Table 3 on individual stocks, but now with the univariate SMB regressions. The conclusion is the same: the sign of the $S - B$ coefficient in the univariate regression fails to identify size.

VI. Conclusion

Our study calls attention to the interpretation of portfolio attribution when self-financing portfolios are included in the attribution model. The self-financing SMB returns used in the Fama–French three-factor model have become standard in academic and industry research as a way to attribute portfolio returns to "size." Whether it is interpreted as a risk factor or capital asset pricing model anomaly, a positive SMB coefficient has been interpreted as evidence of exposure to small-sized companies. In the three-factor regressions, however, the dominance of large-cap stocks in the cap-weighted market factor means that large-cap portfolios can have positive SMB coefficients. Our study also shows that the tendency for a positive factor loading on SMB holds for large-cap individual stocks. Hence, in spite of what it looks like at first glance, $\beta_{SMB} > 0$ does not mean small.

References

Bassett, G. W., and H.-L. Chen, 2001, Portfolio style: Return-based attribution using quantile regression, *Empirical Economics* 26, 293–305.

Carhart, M. M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chan, L. K. C., H.-L. Chen, and J. Lakonishok, 2002, On mutual fund investment styles, *Review of Financial Studies* 15, 1407–37.

Daniel, K., and S. Titman, 1997, Evidence on the characteristic of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Elton, E. J., M. J. Gruber, and C. R. Blake, 2011, Holdings data, security returns, and the selection of superior mutual funds, *Journal of Financial and Quantitative Analysis* 46, 341–67.

Fama, E. F., and K. R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, E. F., and K. R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, E. F., and K. R. French, 2010, Luck versus skill in the cross-section of mutual fund returns, *Journal of Finance* 65, 1915–47.

Korkie, B., and H. J. Turtle, 2002, A mean-variance analysis of self-financing portfolios, *Management Science* 48, 427–43.

Sharpe, W. F., 1992, Asset allocation: Management style and performance measurement, *Journal of Portfolio Management* 18, 7–19.